Measuring the Importance of Learning from Instruction

The Outcome-Consequence Model of Evaluation

Michael J. Hannafin *Assistant Professor University of Colorado Boulder, CO 80309*

Abstract. Traditional models for the evaluation of instructional products and systems tend to be unnecessarily narrow. Often, the importance of learning beyond the immediate instructional setting is either not emphasized or completely overlooked. The Outcome-Consequence Model (OCM) emphasizes the importance of evaluating instructional products and systems using a broadened perspective of the intent of learning from instruction. Specifically, OCM emphasizes evaluation procedures that are systematic and empirical, establish legitimate and appropriate bases for comparison of instructional effectiveness, and evaluate the consequences of learning beyond the immediate temporal environment. While traditional evaluation models may be simpler, the consequences of learning are often not considered when using such models. Instructional designers and evaluators need to direct more attention to the *importance* of learning, and not simply the *act* of learning from instruction.

Evaluation of instructional products and systems has been the focal point of considerable study during the past 15 years (Alkin, Kosecoff, Fitz-Gibbon, & Seligman, 1974; Davis, Alexander, & Yelon, 1974; Hively, Maxwell, Rabehl, Sension, & Lundin, 1973; Worthen, 1977). Generally, the instructional development profession appears to be in substantial agreement regarding the broad classifications of the types of evaluation, formative and summative, applied to typical instructional development projects (Dick 1977a, 1977b; Kaufman, 1972; Gagne & Briggs, 1979; Briggs & Wager, 1981). However, questions regarding the value of much traditional evaluation activity have been raised (Glass, 1976; Guba, 1978; Hannafin, in press; House, 1976). The problem, it appears, is related not to the broad classification of evaluation, but to the specific conceptual validity of evaluation questions raised, the manner in which such questions are operationalized, and the limited scope with which instructional products and systems have been evaluated.

The purposes of this paper are:

(1) to expand the notion of evaluation to include both intuitive and empirical planning considerations;

(2) to examine the "tunnel vision" effects of simple outcome-based evaluation; and

(3) to propose an empirical evaluation model, the Outcome-Consequence Model, that includes not only the immediate outcomes of instruction, but the consequences of instruction as well.

The Definition of Evaluation Revisited

Evaluation has been defined and redefined continually throughout the past two decades (see, for example, Ebel, 1965; Gagne & Briggs, 1979; Popham, 1972; Provus, 1969; Stake, 1967; Stufflebeam, Foley, Gephart, Guba, Hammond, Merriman, & Provus, 1971; Thorndike & Hagen, 1969). One of the least cluttered definitions of evaluation has been offered by Worthen and Sanders (1973), who define it simply as "...the determination of the worth of a thing" (p. 19). It is precisely the determination of the "worth" of instructional programs that has led to a broadened interpretation of the intent and purpose of evaluation. Worth, from an instructional standpoint, may be perceived as acquiring a skill, changing an attitude, or any of a variety of other learning outcomes. It has seldom, however, been interpreted within the context of the *value*, in either a qualitative or quantitative sense, of learning outcomes beyond an immediate instructional setting. It is for this reason that the Outcome-Consequence Model (OCM) of instructional product evaluation has been devised.

Outcome-Consequence Model (OCM) of Instructional Product Evaluation

The Outcome-Consequence Model (OCM) is a systematic procedure which is used to evaluate both current learning outcomes *and* the implications of these outcomes across time. Outcomes may be considered as the acquisition of an attitude, performance of a skill, or other learning outcomes of interest. The rationale underlying OCM is that instructional effects are best considered within a context that permits the evaluator to examine the *consequences* of learning, given past, present, and future influences. In order to examine such consequences, the relationships among antecedent, current, and future outcomes must be systematically identified and operationalized. The OCM evaluator:

(1) establishes a systematic and empirical approach to the examination of evaluation questions of interest;

(2) establishes a legitimate basis for performance comparison;

(3) considers the consequences of learning beyond the immediate instructional setting. The rationale, purpose, and characteristics of each OCM component are outlined in Table 1, and described below in detail.

Systematic and Empirical

Traditional evaluation is often so simplified that highly relevant questions regarding representativeness of groups, performance variance, and the effects of learner characteristics on product effectiveness are not considered. In OCM, the manner in which the intuitive

TABLE 1
Rationale, Purpose, and Characteristics of Each Outcome Consequence Model (OCM) Component

Component	Rationale	Purpose	Characteristics
1.0 Systematic and Empirical	Evaluation should be an orderly, systematic, and operationalized plan of both underlying assumptions and resulting products.	To provide a pre-formulated, empirical set of anticipated outcomes, including both performance-based and attitudinal considerations, in quasi-hypothesis form.	Anticipated outcomes—outcome-based attitudes and interactions are specified in a causal manner to locate more accurately the "real" sources of effects.
2.0 Basis for Comparison	Evaluation should reflect meaningful and appropriate comparisons among legitimate instructional objectives, not simply performance vs. non-performance.	To determine, for informed decision-making purposes, the actual increments in obtained performance compared with pre-existing or readily available alternatives.	Obtained outcomes are compared with assumed zero-level performance, existing knowledge levels (e.g., pretest, performance in previous instructional units, etc.), and expected levels of performance based on existing instruction.
3.0 Temporal Considerations	Consequences and meaning of current learning are best understood when the effects of prerequisite skills on current skills and current skills on subsequent, related skills are systematically evaluated.	To systematically evaluate the effects of different levels of supposed prerequisite skills on current performance, and to determine the carry-over effects of current learning on subsequent, related learning.	Reflects the reality of instructional installation, and projects the consequences of the existence, or non-existence, of prerequisite skills. Also, the extent to which current learning effects subsequent learning under similar and dissimilar instruction is examined.
(a) Past Influences	Past, related, learning exerts a moderating influence on current learning—merely specifying supposed prerequisite skills neither establishes the extent to which such skills are actually needed, nor reflects the reality of typical instructional product or system installation.	To evaluate empirically the extent to which current learning is likely to be moderated by the presence or absence of prespecified prerequisite skills.	Prerequisite skills—their presence, degree of presence, or absence—are "blocked" by objectives in matrix form. Judgments regarding the extent to which instruction can be provided with varying degrees of prerequisite skills are made.
(b) Sustained Effects	Performance effects resulting from product/system installation should be sustained even when subsequent instruction is provided under traditional models.	To evaluate the extent to which performance carry-over from current instruction affects the learning of subsequent, related information presented under traditional instruction.	The consequences of implementing novel instruction within traditional instruction is examined. Also, projected sustained differences—a goal of instruction—are systematically verified.
(c) Cumulative Effects	Performance under several, related units of sequenced instruction—all under a new instructional format—should accumulate compared with either no new instruction or partial new instruction.	To evaluate the extent to which desired performance accumulates as a function of the amount of sequenced instruction delivered under new vs. old formats.	Expands traditional notions of product or system verification to include hypothesized effects of systematic "growth" under new instruction.

judgments of instructional evaluators are operationalized is critical. Using hypothesis-like statements, the OCM evaluator converts intuitive or assumed beliefs regarding the instruction and associated effects into direct statements of anticipated outcomes. For example, the evaluation of a computational mathematics program, designed for low achieving math students, might be stated as "Concept A attainment for low achieving students is equal to or is greater than average achieving students under existing instruction." In effect, the logic that guided design decisions is evaluated as well as the actual products and procedures. Learning or failing to learn Concept A, in this example, is formalized not as a single entity, but as a function of the assumptions and design considerations affecting the instructional system.

Basis for Comparison

Designers of instruction interpret evaluation findings as indications of product effectiveness. Often, this is accomplished by assessing individual or group performance on instructional objectives, with the criterion for effectiveness expressed as some percentage of correct learner responses. This is a sufficiently straightforward process, and quite appropriate for completely novel ID projects. In many cases, however, the need for instructional development becomes apparent after formal needs assessment finds existing instruction to be inadequate, i.e., learning has occurred under existing instruction, but not to satisfactory levels. In such cases, two bases of comparison are appropriate:

(1) the obtained versus desired outcomes as typically determined, and

(2) the difference, positive or negative, between the outcomes produced by the new versus the existing instruction.

When desired outcome levels have been attained, it may be of interest, but of little critical value, to determine formally the performance difference between new and old instruction. However, when desired levels have not been attained, the comparative value of new-versus-old instruction becomes important. In the previously described example of the mathematics program evaluation, the Concept A attainment of low achieving students may have reached the level of desired attainment. However, when compared with concept attainment under existing instruction, the newly developed systems product may be far inferior. Unless an appropriate basis for comparison is established before-the-fact, conclusions regarding product effectiveness must be drawn based on a simplified learning-by-objective paradigm or on the subjective judgments of decision makers.

Temporal Considerations

Traditional evaluation models tend to treat skills as mutually exclusive islands of knowledge, placing little emphasis on the way prior learning influences current learning, or current learning influences subsequent learning. This is, perhaps, the point of greatest departure between traditional and OCM evaluation models. When using OCM evaluation, a designer will systematically investigate how different levels of prerequisite skills affect new skill acquisition, and provide a way to evaluate how well new skills are sustained, and how the newly acquired skills affect the learning of subsequent skills. Instructional evaluation is broadened to emphasize the practical implications and consequences of instruction in a more realistic temporal manner.

Past influences. Virtually all instructional designers recognize that existing levels of knowledge influence the learning of subsequent, related information. In many cases, this recognition leads designers to specify prerequisite skills. While this is likely the most expedient approach, it is seldom possible to limit instruction only to those learners who meet rigidly defined prerequisite skill requirements—especially when designing large projects. Typically, instructional evaluation plans do not account effectively for the different levels of background skills likely to be encountered in a widely used instructional program.

In OCM, by contrast, use of assumed prerequisite skills can provide a fairly sophisticated multi-dimensional perspective to the evaluation of instructional program effectiveness. OCM considers both the *ideals* of product use (i.e., mastery of prerequisite skills), as well as the *realities* of product use (i.e., different levels of prerequisite skills), and the way both cases influence learning.

In the design, development, and evaluation of a large-scale mathematics system, for example, it is often useful to determine the legitimacy of assumed prerequisite skills in moderating related concept acquisition. In the case of computation of simple interest, the presence or absence of various prerequisite skills will undoubtedly affect the learner's acquisition of the skill. It will do so, however, in a differential manner. Whole number multiplication, for example, will likely be far more critical than whole number-times-decimals multiplication—a skill that may actually be mastered through well-designed instruction in computation of simple interest. In practice, the OCM evaluator assesses the consequences of the presence, or absence, of salient prerequisite skills on the attainment of criterion information. As a result, the designer obtains realistic knowledge of the effects of different levels of prerequisite skills on criterion learning. Thus, it becomes possible to forecast the relative consequences of providing new instruction when various prerequisite skills are lacking. These procedures are more meaningful than the simple effective-not-effective verdict, or than the rigid requirement of specified prerequisite skills.

Sustained effects. Information not only has to be learned; it also must be retained. A novel style of instruction might result in the initial acquisition of desired skills, but little retention and subsequent application of these skills (cf. Gay, 1976). Knowing this, evaluators often plan follow-up studies to verify the retention of learned skills. While this is useful, it is inadequate in many instances. Consider, for example, a case where instruction was designed to "fill a hole" in an otherwise adequate instructional sequence. The instructional designer and OCM evaluator must be concerned not only with isolated acquisition and retention of the "missing skill," but also with the transfer of that skill within the larger, intact system.

It is possible to design effective within-skill instruction that is structurally and procedurally incompatible with the existing program. For this reason, evaluation should include an indication of the cross-skill compatibility and transfer among instructional units. OCM emphasizes sustained effects in order to evaluate the extent to which initial learning has been subsequently retained, and the extent to which initial learning under the new system has affected subsequent learning under existing instructional systems of different philosophical, structural, and procedural formats.

Cumulative effects. Existing evaluation designs often include fairly detailed cross-tabulations of the performance of learners by objective, but seldom include a growth component. In OCM, the intent of instruction is evaluated according to anticipated cumulative effects as well as the performance-by-objective analysis used in most traditional evaluation plans. Mathematics learning, for example, tends to be hierarchical in nature. Presumably, student learning in a comprehensive mathematics program should be progressively greater as the number of program-based instructional units is increased, i.e., learning difference effects should accumulate compared with existing instruction and not simply elevate at a single point and maintain only the initial constant difference. In traditional evaluation, skill transfer is assumed but not tested; in OCM, isolated outcomes alone are viewed as insufficient evidence of assumed cumulative effects. In OCM, the consequences of instruction are verified, not simply assumed.

Instructional designers must make certain that *someone* within the development team assumes responsibility for consideration of the meaning of instructional effectiveness beyond the immediate instructional setting.

Teams as Designs: Territoriality in Evaluation

Carey and Briggs (1977) emphasized the importance of interactive teams in instructional development. During recent years, however, a predictable territorial debate over important ID issues has often surfaced among instructional designers, traditional curriculum planners, media producers, and evaluators. Curriculum specialists, for example, have asserted that many of the temporal considerations advocated in OCM fall in the domain of curriculum planning, and not instructional design. Similarly, evaluators often perceive these considerations as evaluation-specific.

Ideally, ID projects should include appropriately placed and qualified personnel to permit role specialization. Under circumstances where such personnel are available, *and* where instructional development is managed effectively, *and* where important relational hierarchies have been identified, *and* where appropriate data exists regarding such skill hierarchies, *and* where curriculum and ID personnel work effectively as a team, the partitioning of responsibilities seems desirable. In many cases, however, ideal conditions simply do not exist. Therefore, instructional designers must make certain that *someone* within the development team assumes responsibility for consideration of the meaning of instructional effectiveness beyond the immediate instructional setting.

Influencing Instructional Design via OCM

It has been suggested that evaluation concerns may be more valuable during the formulation phase of an instructional development project than during the evaluation phase (Morris & Fitz-Gibbon, 1978; Hannafin, in press). It makes little sense to develop a sophisticated evaluation model to examine issues of interest when such issues were not considered initially during the formulation and production phases of development—effects are not likely to be found. On the other hand, careful consideration of the consequences of instruction and learning, applied during the early stages of development, is likely to provide instructional designers with additional relevant product specifications. When OCM considerations are made before-the-fact, instructional designers are more likely to be alert to issues such as across-skill instructional compatibility. In this regard, OCM is valuable in that an expanded formalization of instructional intent is articulated before-the-fact. Table 2 contains a series of questions typically applied during both the development and evaluation phases of each component of OCM.

While valuable during formulation and evaluation, such considerations seem most fruitful when an instructional program is first conceptualized and designed. Certainly, if some things are not considered early in the design of a project, many of the sources of data necessary for answering OCM questions will be either unavailable or difficult to collect. It seems at least as reasonable to use OCM as a guide to instructional planning as to use OCM as a procedure for instructional research.

Closing Comments

It should be apparent to the reader that OCM, as a procedure, will not apply in its entirety to all instructional system and product evaluation activities. However, it should also be apparent that the *rationale* underlying OCM is generalizable to virtually all instructional development and evaluation. Broadened planning, to include systematic consideration of the outcomes *and* consequences of instructional systems and products, is likely to improve the quality and meaningfulness of both development and evaluation.

OCM is not a new classification of evaluation; it is compatible with several existing evaluation models. Instead, the rationale and emphasis of OCM provide a broadened perspective on the process and purpose of instructional evaluation. The unique emphasis of OCM, however, requires instructional systems designers, developers, and evaluators to adopt an expanded perspective on the basic reasons for developing instructional programs.

For years, educators—instructional designers in particular—have been criticized for encouraging "tunnel vision" learning through the extensive use of objectives in materials design, implementation, assessment of student learning, and evaluation of the effectiveness of instructional products (Eisner, 1971; Meyen, 1976; Macdonald-Ross, 1973). On one hand, instructional developers argue that objectives, as such, do not necessarily foster restrictive learning (Popham, 1971). On the other hand, developers make distressingly few attempts to systematically disprove the criticism. The "big picture" of instruction—intent, effectiveness, consequence, influence—must be reconciled. As a profession, developers must not disregard the broader issues of instructional development. While the OCM model of evaluation is a modest step

Sample Considerations in Planning for Each Component of OCM

1.0 Systematic and Empirical	2.0 Basis for Comparison	3.0 Temporal Considerations: (a) Past Influences	(b) Sustained Effects	(c) Cumulative Effects
What are the specific hypotheses for outcomes of each portion of instructional product or system? What are the projected relationships among outcomes? What data will be needed to address each hypothesis? What are the assumptions underlying the design of instruction? What type of analysis, if any, will account best for measurement error?	What is the performance history of learners in area of instruction? Will a pretest be administered? What are the specific objectives for which instruction has been designed? What legitimate instructional alternatives exist? What instruction, if any, has been used to date?	What are the assumed prerequisite skills? What is the effect of the absence or presence of such skills on current skill acquisition? Which assumed prerequisite skills are most critical to the acquisition of current content? Can critical prerequisite skills instruction be embedded into new instruction?	For which subsequent instruction will current skills be most highly dependent? To what extent will the acquisition of skills via new instruction provide a continued advantage during subsequent learning? Does the variation in instructional format from new to existing instruction cause interference during the switching? Will transitional instruction, designed to blend new and old formats be helpful or necessary?	How many sequential instructional products/units/modules are to be developed? Which learners will participate in different phases of instruction? During tryouts, can learners be assigned to different combinations of new instruction, ranging from no new instruction to all new instruction? Will learning of prescribed content be affected positively by the *amount* of new-versus-old instruction?

toward this end, it represents a recognition of the important consequences of systematic instructional development.

References

Alkin, M., Kosecoff, J., Fitz-Gibbon, C., & Seligman, R. *Evaluation and decision making: The Title VII experience*. Los Angeles: Center for the Study of Evaluation, 1974.

Briggs, L., & Wager, W. *Handbook of procedures for the design of instruction* (2nd ed.). Englewood Cliffs, N.J.: Educational Technology Publications, 1981.

Carey, J. O., & Briggs, L. Teams as designers. In L. Briggs (Ed.), *Instructional design—Principles and applications*. Englewood Cliffs, N.J.: Educational Technology Publications, 1977.

Davis, R., Alexander, L., & Yelon, S. *Learning system design—An approach to the improvement of instruction*. New York: McGraw-Hill, 1974.

Dick, W. Formative evaluation. In L. Briggs (Ed.), *Instructional design—Principles and applications*. Englewood Cliffs, N.J.: Educational Technology Publications, 1977. (a).

Dick, W. Summative evaluation. In L. Briggs (Ed.), *Instructional design—Principles and applications*. Englewood Cliffs, N.J.: Educational Technology Publications, 1977. (b).

Ebel, R. *Measuring educational achievement*. Englewood Cliffs, N.J.: Prentice-Hall, 1965.

Eisner, E. Educational objectives: Help or hindrance? In M.B. Kapfner (Ed.), *Behavioral objectives in curriculum development*. Englewood Cliffs, N.J.: Educational Technology Publications, 1971.

Gagne, R., & Briggs, L. *Principles of instructional design* (2nd ed.). New York: Holt, Rinehart and Winston, 1979.

Gay, L. *Educational research*. Columbus, OH: Charles E. Merrill Publishing Company, 1976.

Glass, G. (Ed.) *Evaluation Studies Review Annual* (Vol. 1). Beverly Hills, CA: Sage Publications, 1976.

Guba, E. *Toward a methodology of naturalistic inquiry in educational evaluation*. Los Angeles: Center for the Study of Evaluation, 1978.

Hannafin, M. Practical evaluation in educational programs. In D. Wiseman (Ed.), *Procedures for planning and developing child service development centers*. Tempe, AZ: Arizona Educational Information System, in press.

Higgins, N., & Sullivan, H. *Writing worthwhile objectives*. Tempe, AZ: Teaching for Competence, Publisher, 1978.

Hively, W., Maxwell, G., Rabehl, G., Sension, D., & Lundin, S. *Domain-referenced evaluation: A technical handbook and a case study from the Minnemast Project*. Los Angeles: Center for the Study of Evaluation, 1973.

House, E. R. Justice in evaluation. In G. V. Glass (Ed.), *Evaluation Studies Review Annual* (Vol. 1). Beverly Hills, CA: Sage Publications, 1976.

Kaufman, R. *Education system planning*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1972.

MacDonald-Ross, M. Behavioral objectives: A critical review. *Instructional Science*, 1973, *2*, 1-52.

Meyen, E. *Developing instructional units* (2nd ed.). Dubuque, IA: Wm. C. Brown Company, Publishers, 1976.

Morris, L., & Fitz-Gibbon, C. *Evaluators Handbook*. Los Angeles: Sage Publications, 1978.

Popham, W. *An evaluation guidebook*. Los Angeles, CA: The Instructional Objectives Exchange, 1972.

Popham, W. Probing the validity arguments against behavioral objectives. In M.B. Kapfner (Ed.), *Behavioral objectives in curriculum development*. Englewood Cliffs, N.J.: Educational Technology Publications, 1971.

Provus, M. Evaluation of on-going programs in the public school system. In R. W. Tyler (Ed.), *Educational evaluation: New roles, new means*. The 68th Yearbook of the National Society for the Study of Education, Part II. Chicago, IL: National Society for the Study of Education, 1969.

Stake, R. The countenance of educational evaluation. *Teachers College Record*, 1967, *68*, 523-540.

Stufflebeam, D., Foley, W., Gephart, W., Guba, E., Hammond, R., Merriman, H., & Provus, M. *Educational evaluation and decision making*. Itasca, IL: Peacock Publishing Company, Inc., 1971.

Thorndike, R., & Hagen, E. *Measurement and evaluation in education and psychology* (3rd ed.). New York, NY: John Wiley, 1969.

Worthen, B. Characteristics of good evaluation studies. *Journal of Research and Development in Education*, 1977, *10*, 3-20.

Worthen, B., & Sanders, J. *Educational evaluation: Theory and practice*. Belmont, CA: Wadsworth Publishing Co., Inc., 1973.